

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY REPORT ON PATENTABILITY (Chapter II of the Patent Cooperation Treaty)

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference 51242WO	FOR FURTHER ACTION See Form PCT/IPEA/416	
International application No. PCT/IB 2002/004990	International filing date (<i>day/month/year</i>) 28-11-2002	Priority date (<i>day/month/year</i>) -
International Patent Classification (IPC) or national classification and IPC G06K 9/00		
Applicant Nokia Corporation et al		

1. This report is the international preliminary examination report, established by this International Preliminary Examining Authority under Article 35 and transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 5 sheets, including this cover sheet.

3. This report is also accompanied by ANNEXES, comprising:

a. (*sent to the applicant and to the International Bureau*) a total of 3 sheets, as follows:

sheets of the description, claims and/or drawings which have been amended and are the basis of this report and/or sheets containing rectifications authorized by this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions).

sheets which supersede earlier sheets, but which this Authority considers contain an amendment that goes beyond the disclosure in the international application as filed, as indicated in item 4 of Box No. I and the Supplemental Box.

b. (*sent to the International Bureau only*) a total of (indicate type and number of electronic carrier(s)) _____, containing a sequence listing and/or tables related thereto, in computer readable form only, as indicated in the Supplemental Box Relating to Sequence Listing (see Section 802 of the Administrative Instructions).

4. This report contains indications relating to the following items:

<input checked="" type="checkbox"/>	Box No. I	Basis of the report
<input type="checkbox"/>	Box No. II	Priority
<input type="checkbox"/>	Box No. III	Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
<input type="checkbox"/>	Box No. IV	Lack of unity of invention
<input checked="" type="checkbox"/>	Box No. V	Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
<input type="checkbox"/>	Box No. VI	Certain documents cited
<input type="checkbox"/>	Box No. VII	Certain defects in the international application
<input type="checkbox"/>	Box No. VIII	Certain observations on the international application

Date of submission of the demand 15-06-2004	Date of completion of this report 25-02-2005
Name and mailing address of the IPEA/SE Patent- och registreringsverket Box 5055 S-102 42 STOCKHOLM Facsimile No. +46 8 667 72 88	Authorized officer Henrik Andersson/MN Telephone No. +46 8 782 25 00

INTERNATIONAL PRELIMINARY REPORT ON PATENTABILITY

International application No.

PCT/IB 2002/004990

Box No. I Basis of the report

1. With regard to the language, this report is based on the international application in the language in which it was filed, unless otherwise indicated under this item.

This report is based on a translation from the original language into the following language _____, which is the language of a translation furnished for the purposes of:

- international search (under Rules 12.3 and 23.1(b))
 publication of the international application (under Rule 12.4)
 international preliminary examination (under Rules 55.2 and/or 55.3)

2. With regard to the elements of the international application, this report is based on (replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report):

the international application as originally filed/furnished

the description:

pages 1-23 as originally filed/furnished

pages* _____ received by this Authority on _____

pages* _____ received by this Authority on _____

the claims:

pages _____ as originally filed/furnished

pages* _____ as amended (together with any statement) under Article 19

pages* 24-26 received by this Authority on 23-12-2004

pages* _____ received by this Authority on _____

the drawings:

pages 1-5 as originally filed/furnished

pages* _____ received by this Authority on _____

pages* _____ received by this Authority on _____

a sequence listing and/or any related table(s) – see Supplemental Box Relating to Sequence Listing.

3. The amendments have resulted in the cancellation of:

the description, pages _____

the claims, Nos. _____

the drawings, sheets/figs _____

the sequence listing (specify): _____

any table(s) related to the sequence listing (specify): _____

4. This report has been established as if (some of) the amendments annexed to this report and listed below had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).

the description, pages _____

the claims, Nos. _____

the drawings, sheets/figs _____

the sequence listing (specify): _____

any table(s) related to the sequence listing (specify): _____

* If item 4 applies, some or all of those sheets may be marked "superseded."

INTERNATIONAL PRELIMINARY REPORT ON PATENTABILITY

International application No.

PCT/IB 2002/004990

Box No. V Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Claims	<u>1-15</u>	YES
	Claims	_____	NO
Inventive step (IS)	Claims	<u>1-15</u>	YES
	Claims	_____	NO
Industrial applicability (IA)	Claims	<u>1-15</u>	YES
	Claims	_____	NO

2. Citations and explanations (Rule 70.7)**Documents cited in the International Search Report:**

D1: CHENG YANG: "MACS: Music Audio Characteristics Sequence Indexing for Similarity Retrieval". In IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. 2001. 21-24 Oct. New York.

D2: US 5402339 A

D3: EP1093109

D1 discloses a method for matching audio data. Audio data is firstly converted into a string of elements. In order to compare two strings, the method comprises indexing means in order to capture the relative order of the elements included in the string. A matching procedure is then performed; each match contains a tuple (query-offset, matching-offset). A "good" match occurs when the relative order of the elements in the query string and the reference string agrees.

D2 discloses an apparatus for retrieving musical information. A music piece is converted to a string of elements, where each element represents note data. The apparatus also include means for producing position data representing positions at which note data is positioned in the musical information. The apparatus also include means for indexing the musical information by storing an index which indicates the relationship between note data items and position data. The relative order of the note data is considered when matching strings of musical information.

The applied invention is a method, software tool, computer

.../...

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.

Continuation of Box V

program product, computer data signal and electronic device for determining and outputting a similarity measure between two data strings. The similarity measure is based on the number of identical elements in both strings and the relative position of the elements in the strings.

In the response to the written opinion, received 23-12-2004, is it stated, in page 1, that the new claim 1 is based on the original claims 1, 2 and 4. However, in page 4 under chapter V. Novelty is it stated that claim 1 is based on the features of the original claims 1, 2 and 5.

This report is base on the first statement, that the new claim 1 is based on the original claims 1, 2 and 4, because if the new claim 1 is studied that seems to be the correct statement. Consequently, the parts of the response that relates to the second statement, in page 4, were not the subject of this report.

Referring to claims 1-15:

Claim 1-15 refers to a method for determining and outputting a similarity measure between two data strings. Relative positions of the elements in the strings are firstly determined. A matching measure is then determined based on how far the relative positions of elements in the second string match with the relative position of elements in the first string. The method also contains a consecutive numbering of same data entities according to their relative position in accordance with the position label.

D1 discloses a method where the relative positions, in two strings representing musical pieces, are considered in order to measure similarity (see sections 2.3, 2.4 and figure 5). The relative position of matching elements plays a crucial role in the similarity measure, see for example figure 5, where two similarity measures gives 10 matches each, but the top one is considered to be a better match because the relative order of elements in string s is almost persevered in string r.

D2 also discloses an invention where the relative position of elements in a string is considered when determining similarity

.../...

INTERNATIONAL PRELIMINARY REPORT ON PATENTABILITY

International application No.

PCT/IB 2002/004990

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.

Continuation of Box V

between two strings. The relative position is given by an index which indicates relationships between elements in the string and their position (see for example claim 1).

In view of the aforementioned, both D1 and D2 disclose inventions where the relative position of elements is used as a similarity criteria when matching two strings.

Therefore, are the documents cited in the International Search Report representations of prior art. The claimed invention stated in claims 1-15 is not considered to be anticipated by these documents. None of the documents or any relevant combination of them reveals a method for determining and outputting a similarity measure between two data strings, where the method also contains a consecutive numbering of same data entities according to their relative position in accordance with the position label, as described in these claims.

According to the arguments stated above, the invention claimed in claims 1-15 is novel, considered to involve an inventive step and have industrial applicability.

24

JC20 Rec'd PCT/PTO 05 MAY 2005

Claims

1. Method for determining and outputting a similarity measure between two data strings, each data string comprising data entities, comprising:
 - receiving a first data string,
 - receiving a second data string,
characterized by
 - determining pairs of consecutively following data entities in said first data string,
 - determining the relative positions of said pairs of consecutively following data entities in said first data string,
 - allocating a position label to each of said data entities in the first data string,
 - numbering same data entities according to their relative position in accordance with the position label,
 - determining similar data entities with the same order in said second data string,
 - determining the relative positions of said determined data entities in said second data string,
 - determining a matching measure by determining how far the relative positions of data entities in said second data string match with the relative positions of consecutively following data entities in said first data string, and
 - outputting a similarity measure which corresponds to the matching measure of at least one comparison result.
2. Method according to claim 1, further comprising:
 - determining at least one error limit for at least one of said entities,
 - considering said at least one error limit during said determination of said matching measure.
3. Method according to claim 1, further comprising:
 - determining a first distance between said two data entities of consecutively following data entities in said first data string,
 - determining a second distance of said two data entities determined in said second data string,
 - determine a difference between said first and second distances, and
 - considering said difference during said determination of said matching measure.

4. Method according to claim 1, further comprising:
 - storing said second string together with said similarity measure.
5. Method according to claim 1, further comprising:
 - determining a threshold for said similarity measure, and
 - outputting said second string, if said determined similarity measure at least equals said threshold.
6. Method according to claim 5, further comprising:
 - repeating said determination of said similarity measure with a number of second strings, and
 - determining said threshold in correspondence with a number of second strings to be outputted.
7. Method according to claim 1, further comprising:
 - analyzing the first string for entities not present in the first string, and
 - suppressing in the second string all said entities not present in said first string.
8. Method according to claim 7, further comprising:
 - determining the number of entities that are present in the second string, but are not present in the first string, as a second similarity measure.
9. Method according to claim 8, further comprising:
 - determining a section within said second string comprising at least the same number of entities that are simultaneously present in both strings.
10. Software tool comprising program code means stored on a computer readable medium for carrying out the method of anyone of claims 1 to 9 when said software tool is run on a computer or network device.
11. Computer program product comprising program code means stored on a computer readable medium for carrying out the method of anyone of claims 1 to 9 when said program product is run on a computer or network device.
12. Computer program product comprising program code, downloadable from a server for carrying out the method of anyone of claims 1 to 9 when said program product is run on a computer or network device.
13. Computer data signal embodied in a carrier wave and representing a program that

23-12-2004

26

instructs a computer to perform the steps of the method of anyone of claims 1 to 9.

14. Electronic device for determining and outputting a similarity measure between two data strings each comprising data entities, comprising:
 - a component for receiving a first data string of entities and a second data string of entities,
 - a processing unit being connected to said receiving component, said processing unit being configured to determine pairs of consecutively following data entities in said first data string, said processing unit being configured to determine the relative positions of said pairs of consecutively following data entities in said first data string, and for allocating a position label to each of said data entities in the first data string, and numbering same data entities according to their relative position in accordance with the position label; said processing unit being configured to determine similar data entities with the same order in said second data string, and to determine the relative positions of said determined data entities in said second data string, , , said processing unit being configured to determine a matching measure by determining how far the relative positions of data entities in said second data string match with the relative positions of consecutively following data entities in said first data string, and said processing unit being configured to output a similarity measure which corresponds to the matching measure of at least one comparison result, and
 - an interface being connected to said processing unit for outputting said similarity measure.
15. Electronic device according to claim 14, further comprising a storage connected to said processing unit for storing received strings and said determined similarity measures.